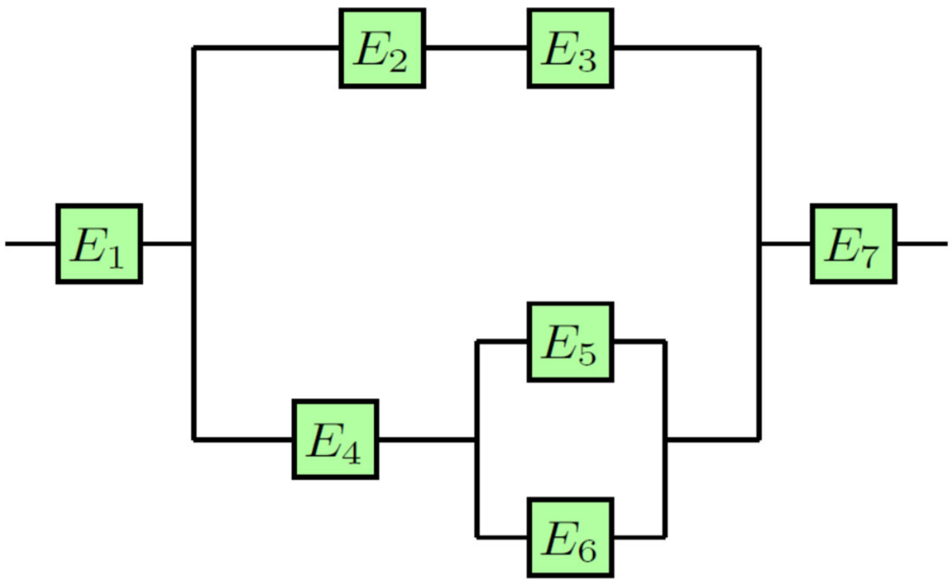
Matlab group exercise

- Test our result for this circuit.
- Download circuit_template.m from the website

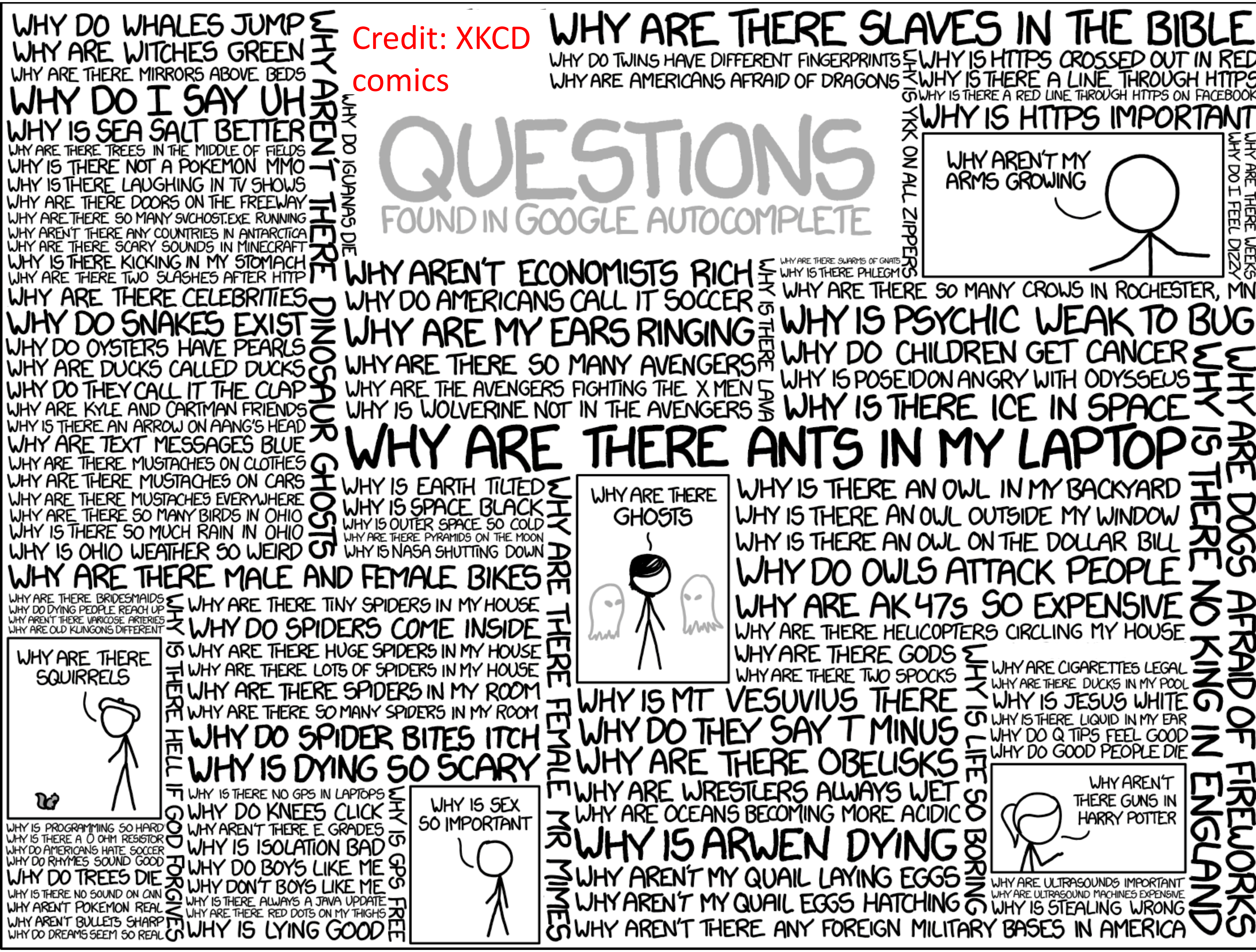


Component	E_1	E_2	E_3	E_4	E_5	E_6	E_7
Probability of functioning well	0.9	0.5	0.3	0.1	0.4	0.5	0.8

P(Works) = 0.9.*(1-(1-0.5.*0.3).*(1-0.1.*(1-0.6.*0.5))).*0.8=0.15084

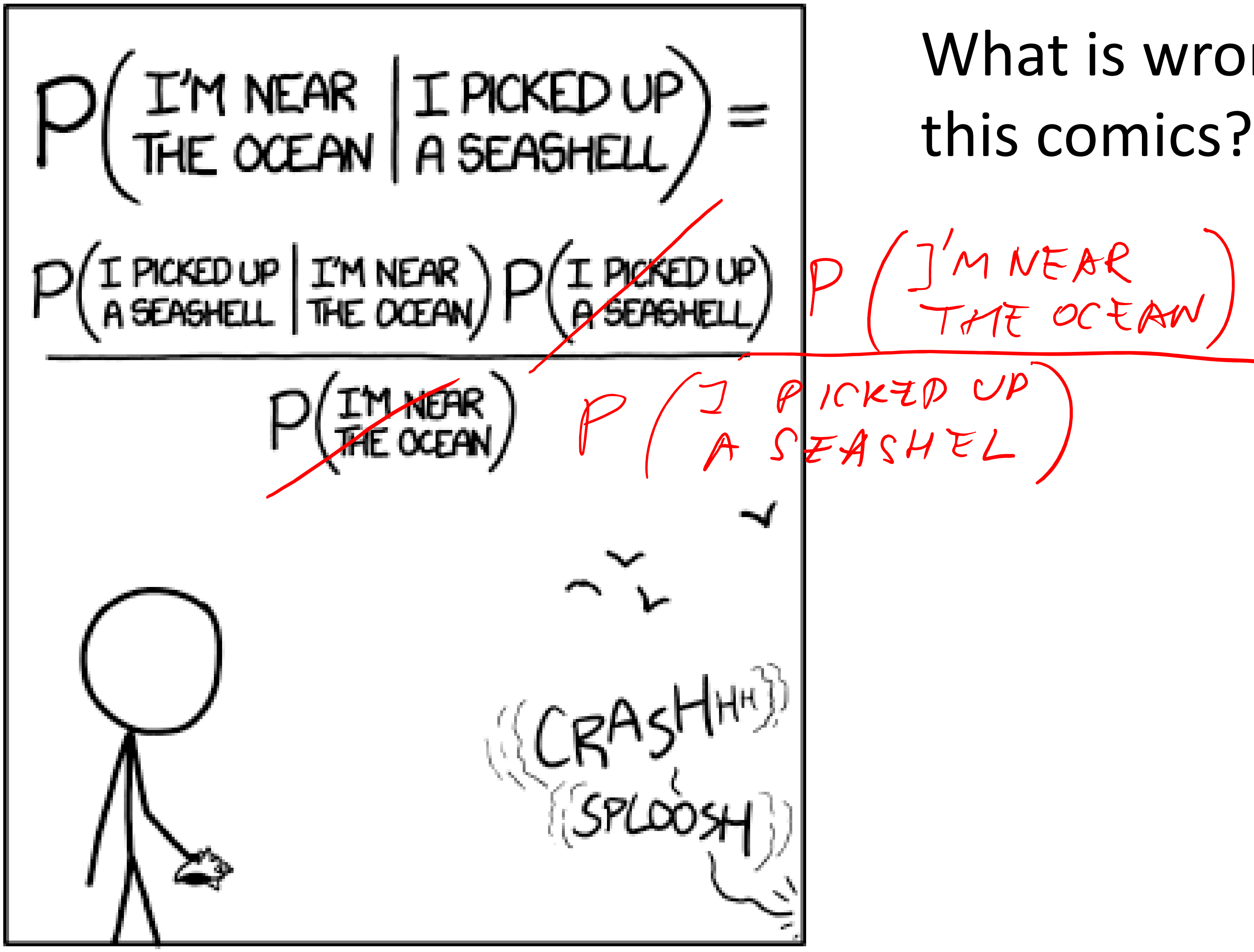
Here is how I did it

- Stats=1e6;
- count= 0;
- for i = 1: Stats
- e1 = rand < 0.9; e2 = rand < 0.5; e3 = rand < 0.3;
- e4 = rand < 0.1; e5 = rand < 0.4; e6 = rand < 0.5;
- e7 = rand < 0.8;
- s1 = min(e2,e3); % or s1 = e2*e3;
- s2 = max(e5,e6); % or s2= e5+e6>0;
- s3 = min(e4,s2); % or s3 = e4*s2;
- s4 = max(s1,s3); % or s4 = s1+s3 > 0;
- s5= min([e1;s4;e7]); % or s5=e1*s4*e7;
- count = count + s5;
- End;
- P_circuit_works = count/Stats
- **% our calculation: P(circuit_works)= 0.9.*(1-(1-0.5.*0.3).*(1-0.1.*(1-0.6.*0.5))).*0.8==0.15084**



Credit: XKCD comics

Reminder:
Conditional probability



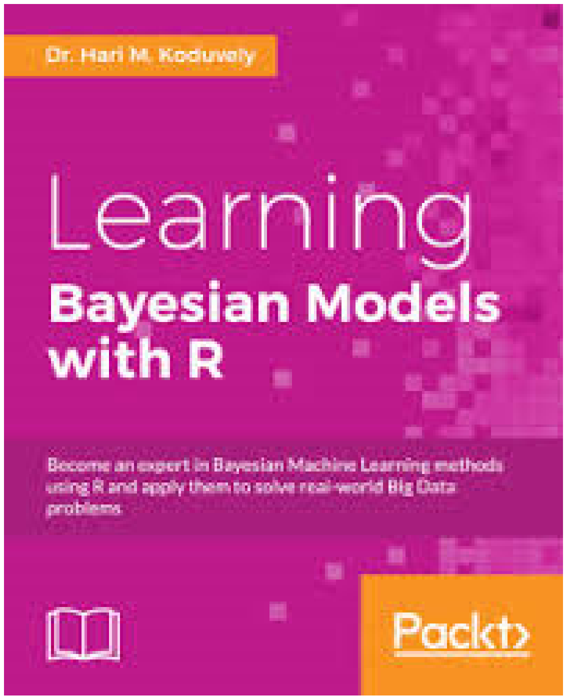
STATISTICALLY SPEAKING, IF YOU PICK UP A SEASHELL AND *DON'T* HOLD IT TO YOUR EAR, YOU CAN PROBABLY HEAR THE OCEAN.

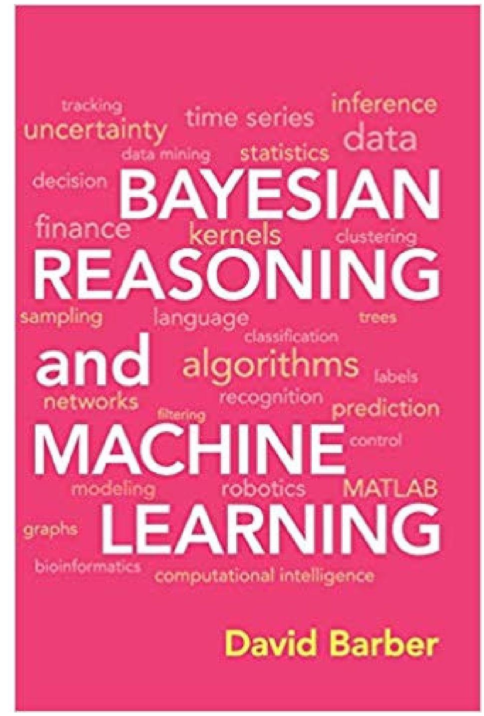
What is wrong in this comics?

If you are not yet reading XKCD comics https://xkcd.com/ you should start

Bayes Theorem

Bayes’ theorem





Thomas Bayes (1701-1761)

English statistician, philosopher, and Presbyterian minister

Bayes’ theorem was presented in "An Essay towards solving a Problem in the Doctrine of Chances" which was read to the Royal Society in 1763 already after Bayes' death.

Bayes' theorem (simple)

$$P(A \cap B) = P(A|B)P(B) = P(B \cap A) = P(B|A)P(A)$$

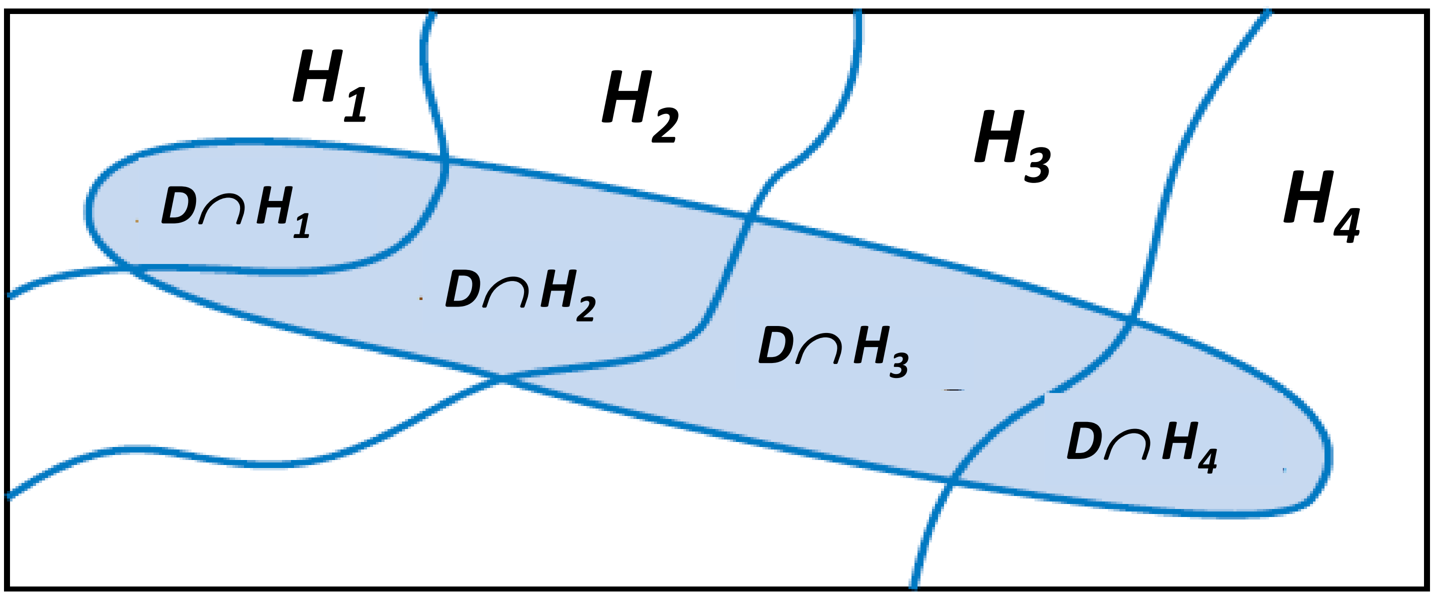
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- In Science we often want to know:
  "How much faith should I put into hypothesis, given the data?"
  or $P(H|D)$ (see also the inductive definition of probability)
- What we usually can calculate if the hypothesis/model is OK:
  "Assuming that this hypothesis is true, what is the probability of the observed data?" or $P(D|H)$
- Bayes' theorem can help: $P(H|D)=P(D|H) \bullet P(H)/P(D)$
- The problem is $P(H)$ (so-called prior) is often not known

Bayes' theorem (continued)

Works best with exhaustive and mutually-exclusive hypotheses:
$H_1, H_2, \ldots H_n$ such that $H_1 \cup H_2 \cup H_3 \ldots \cup H_n = S$ and $H_i \cap H_j = \emptyset$ for $i \neq j$

$P(H_k|D)=P(D|H_k) \bullet P(H_k)/P(D)$
where:
$P(D)= P(D|H_1) \bullet P(H_1) + P(D|H_2) \bullet P(H_2) + \ldots P(D|H_n) \bullet P(H_n)$



An **awesome new test** has been invented for an early detection of cancer. The probability that it correctly identifies someone with cancer as positive is 95%, and the probability that it correctly identifies someone without cancer as negative is 99%. The incidence of this type of cancer in the general population is 10^{-4}. A random person in the population takes the test, and the result is positive.

What is the probability that he/she has cancer?

A. 99%
B. 95%
C. 30%
D. 1%

Get your i-clickers

An **awesome new test** has been invented for an early detection of cancer. The probability that it correctly identifies someone with cancer as positive is 95%, and the probability that it correctly identifies someone without cancer as negative is 99%. The incidence of this type of cancer in the general population is 10^{-4}. A random person in the population takes the test, and the result is positive.

What is the probability that he/she has cancer?

A. 99%
B. 95%
C. 30%
D. 1%

Get your i-clickers

participants

10^6 participants → 100 cancer → 95 positive tests

→ $10^6 - 100 \approx 10^6$ no cancer

10^6 participants with no cancer → 10,000 positive tests

$$P(C|P) = \frac{95}{10{,}000 + 95} \approx 1\%$$

Events: C - cancer, C' - no cancer

Test events Y - positive, N - negative

We know:

$P(C) = 10^{-4}$, $P(Y|C) = 0.95$

$P(N|C') = 0.99$

We need

$P(C|Y)$

Bayes:

$$P(C|Y) = P(Y|C) \cdot \frac{P(C)}{P(Y)}$$?

P(Y) - probability that a random person will test positive

$$P(Y) = P(Y \cap C) + P(Y \cap C') =$$

$$= P(Y|C)P(C) + P(Y|C')P(C') =$$

$$= 0.95 \times 10^{-4} + (1 - 0.99) \times (1 - 10^{-4}) \approx$$

$$\approx 10^{-4} + 10^{-2} \approx 10^{-2} = 1\%$$

$$P(C|Y) = P(Y|C) \cdot \frac{P(C)}{P(Y)} = 0.95 \times \frac{10^{-4}}{10^{-2}} \approx$$

$$\approx 1\%$$

An **awesome new test** has been invented for an early detection of cancer. The probability that it correctly identifies someone with cancer as positive is 95%, and the probability that it correctly identifies someone without cancer as negative is 99%. The incidence of this type of cancer in the general population is 10^{-4}. A suspected cancer patient with likelihood of cancer 50% takes the test, and the result is positive.

What is the probability that he/she has cancer?

A. 99%
B. 95%
C. 30%
D. 1%

Get your i-clickers

An **awesome new test** has been invented for an early detection of cancer. The probability that it correctly identifies someone with cancer as positive is 95%, and the probability that it correctly identifies someone without cancer as negative is 99%. The incidence of this type of cancer in the general population is 10^{-4}. A suspected cancer patient with likelihood of cancer 50% takes the test, and the result is positive.

What is the probability that he/she has cancer?

A. 99%
B. 95%
C. 30%
D. 1%

Get your i-clickers

What if a doctor is already 50% sure of cancer based on other tests?

That changes things!

Now $P(C) = P(C') = 0.5$

$$P(C|Y) = \frac{P(Y|C) \cdot P(C)}{P(Y|C) \cdot P(C) + P(Y|C') P(C')} =$$

$$= \frac{0.95 \times 0.5}{0.95 \times 0.5 + (1 - 0.99) \times 0.5} \approx 0.99$$

How come?
I thought it was a ***great*** test..

- Let C – be the event that the patient has cancer; C' – patient is cancer free
- Y/N – events that test is Positive/Negative (N=Y')
- We know: P(C)=10^{-4}, P(Y|C)=0.95, P(N|C')=0.99
- We need to find P(C|Y)
- Bayes to the rescue: P(C|Y)=P(Y|C)*P(C)/P(Y)
- What on earth is P(Y) ???

What on Earth is P(Y) ???

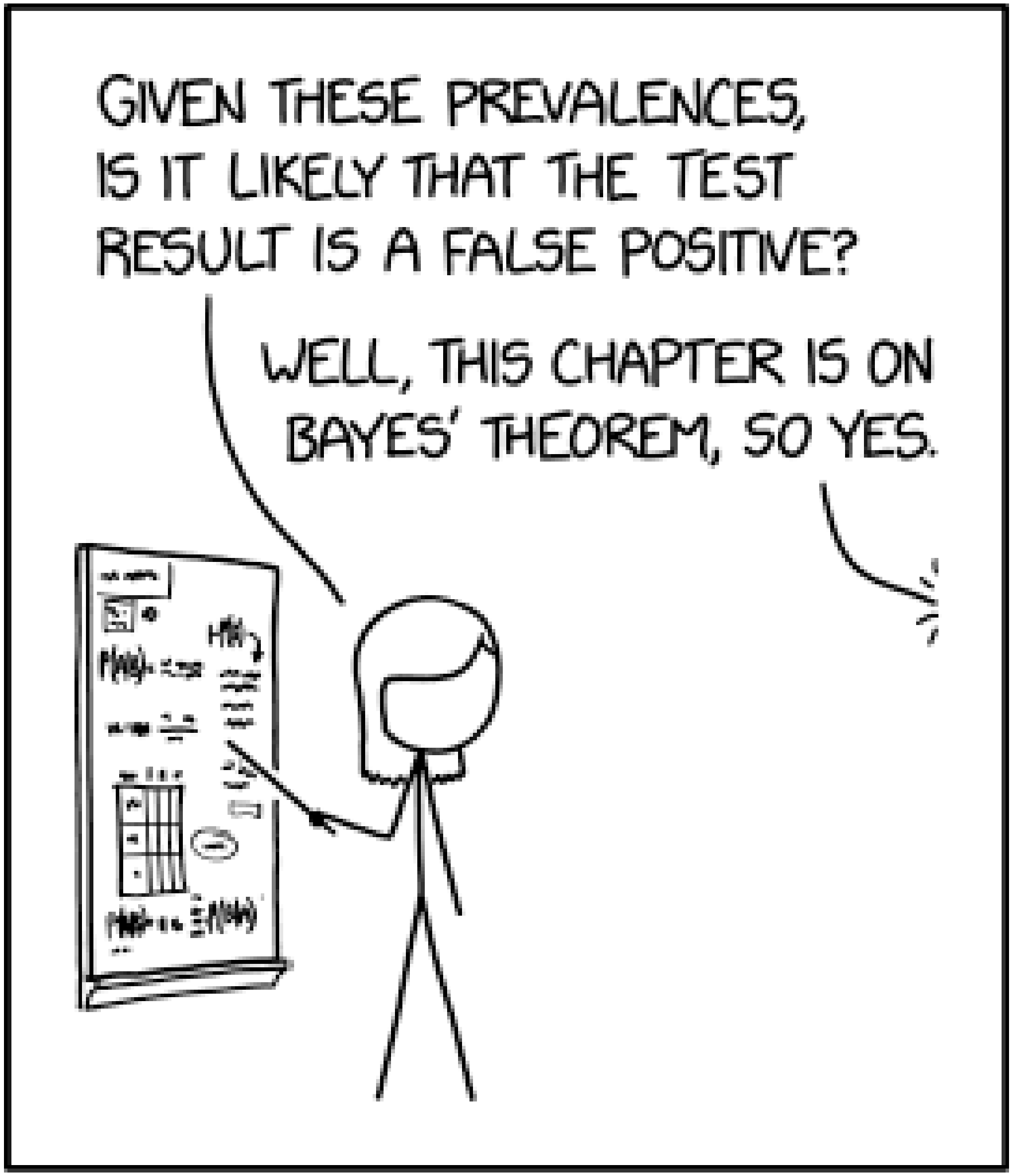
- Likelihood that a random patient would test Y: $P(Y)=P(Y \cap C)+P(Y \cap C')=P(Y|C)P(C)+P(Y|C')P(C')= 0.95*10^{-4}+(1-0.99)*(1-10^{-4})\approx 0.01$
- Hence $P(C|Y)=P(Y|C)*P(C)/P(Y) \approx 10^{-4}/0.01=0.01=1\%$
- But we would like it to be 100%, please!!!
- Until the false positive discovery rate $1-P(N|C')$ does not fall below the general population prevalence the result will never be close 100%

What if I am already 50% sure (based on other tests) that a patient has cancer?

- That changes everything!
- Now P(C)=P(C')=0.5
- $P(C|Y)=P(Y|C)*P(C)/[P(Y|C)*P(C)+ P(Y|C')*P(C')]=$

$0.95*0.5/[0.95*0.5+(1-0.99)*0.5]=0.99$

- Now the doctor can be almost 100% sure.
- The importance of prior:
  - If prior belief that one has cancer is 10^{-4} – test is useless
  - If prior belief is at least 1% - the test is useful



SOMETIMES, IF YOU UNDERSTAND BAYES' THEOREM WELL ENOUGH, YOU DON'T NEED IT.

If you are not yet reading XKCD comics https://xkcd.com/ you should start

Sensitivity/specificity of the standard test for prostate cancer: <u>PSA level > 4.0ng/mL</u>

- Sensitivity of the test is 71.9%
  - fraction of cancer patients who will test positive
  - False negative rate is 28.1%
- Specificity of the test is 90%
  - fraction of healthy patients who will test negative
  - False positive rate is 10%

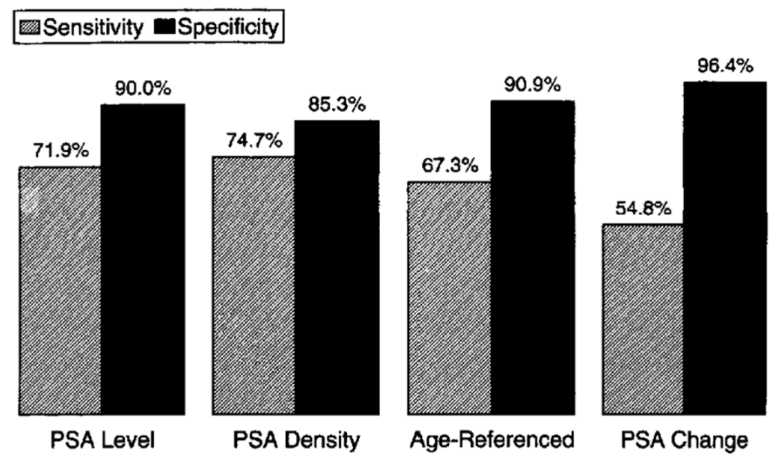


Figure 1. The relative sensitivity and specificity of different indexes of prostate specific antigen (PSA). Except for PSA change, sensitivity is the proportion of 171 known cancers cases for whom the index was positive and specificity is the proportion of 2011 men with normal transrectal ultrasound and digital rectal examinations not known to have prostate cancer who were negative on the index. The sensitivity and specificity of PSA change was evaluated in 84 men with prostate cancer and 1473 men without prostate cancer for on whom multiple PSA measures were available. A PSA level of 4.0 ng/ml or less was considered normal. A PSA density of 0.1 ng/ml per cubic centimeter of ultrasound-measured gland volume was considered normal. Age-referenced PSA was considered normal if it was 3.5 ng/ml or less in men aged 50–59, 4.5 ng/ml in men aged 60–69, and 6.5 ng/ml in men aged 70–79. PSA change was considered normal if the annual rate of PSA change was … or less per year.

Mettlin C, Littrup PJ, Kane RA, Murphy GP, Lee F, Chesley A, et al. Cancer. 1994;74: 1615–1620.

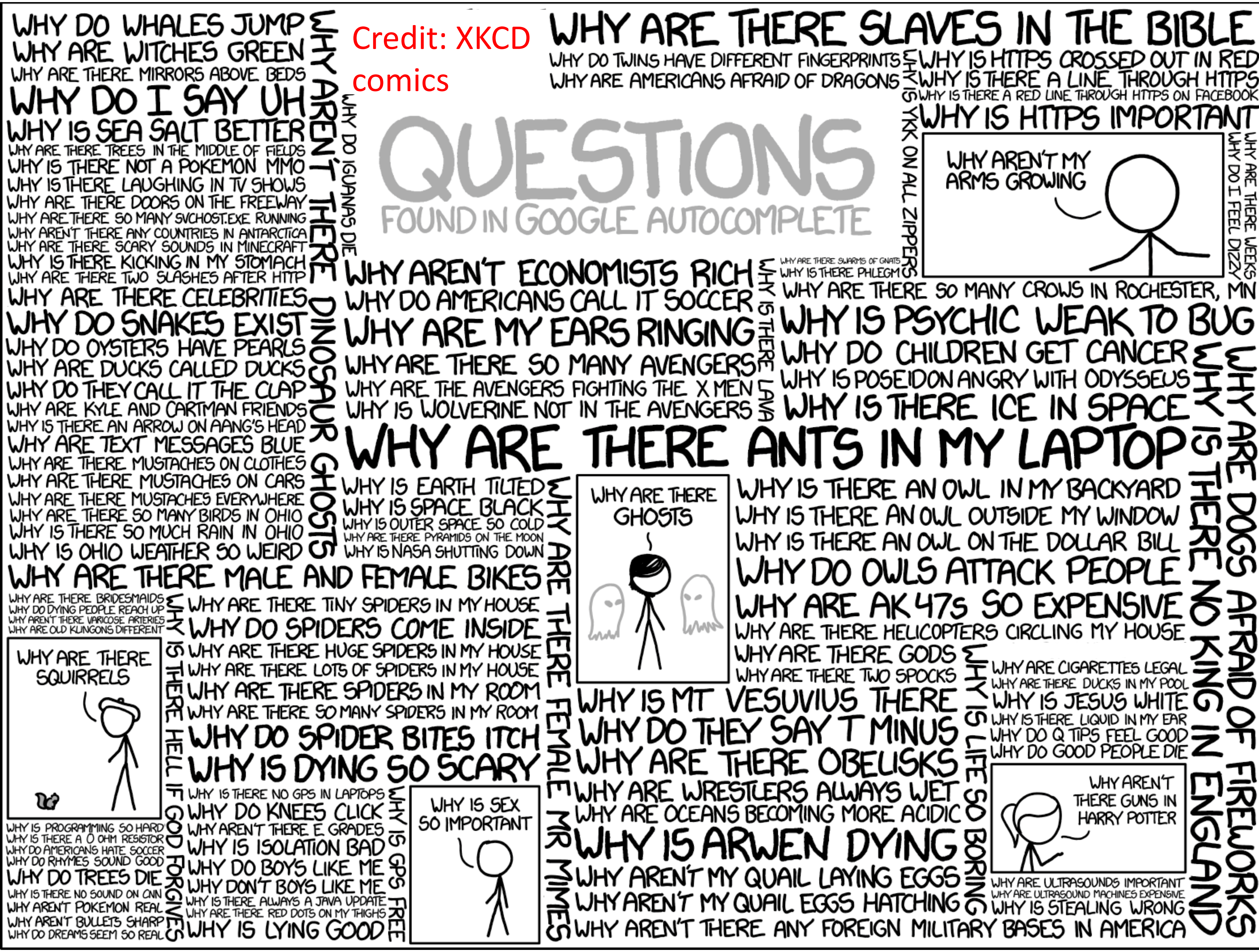
Prostate cancer is the most common type of cancer found in males. It is checked by PSA test that is notoriously unreliable. The probability that a noncancerous man will have an elevated PSA level >4.0 ng/mL is approximately 0.1, with this probability increasing to approximately 0.719 if the man does have prostate cancer. If, based on other factors, a physician is 50 percent certain that a male has prostate cancer, what is the conditional probability that he has the cancer given that the test indicates an elevated PSA level?

A. 99.9%
B. 95%
C. 88%
D. 55%

Get your i-clickers

All this trouble for a lousy 38% gain in confidence? I don't believe you!

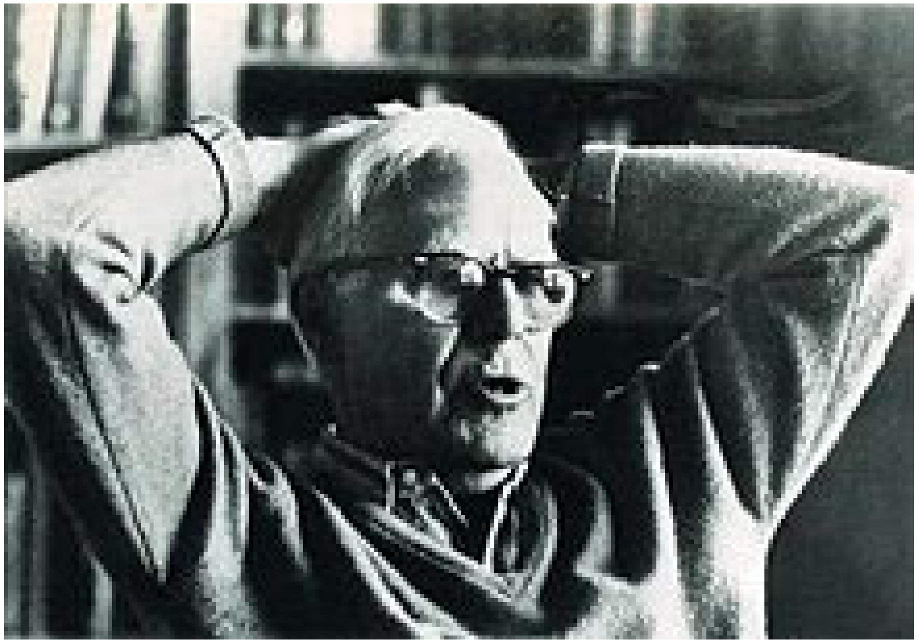
- Let C – be the event that the patient has cancer; C' – patient is cancer free, E – events that the PSA test was elevated
- We know **doctor's prior belief**: P(C)=0.5
- We know test stats: P(E|C)=0.719, P(E|C')=0.1
- We need to find P(C|E)=P(E|C)*P(C)/P(E)
- P(E)=P(E|C)*P(C)+P(E|C')*P(C')= =0.719*0.5+0.1*0.5=0.41
- P(C|E)=0.5*0.719/0.41=0.88 or 88%



Credit: XKCD comics

Secretary problem

- An employer has a known number – n – of applicants for a secretary position, whom are interviewed one at a time
- Employer can easily evaluate and rank applicants relative to each other but has no idea of the overall distribution of their quality
- Employer has only one chance to choose the secretary: gives yes/no answer in the end of each interview and cannot go back to rejected applicants
- How can employer maximize the probability to choose the best secretary among all applicants?



Martin Gardner (1914 – 2010)
Described the "secretary problem"
in Scientific American 1960.
was an American popular mathematics and popular science writer. Best known for "recreational mathematics": He was behind the "Mathematical Games" section in Scientific American.

Eugene Dynkin (1924 – 2014)
solved this problem in 1963.
He referred to it as a "picky bride problem"
was a Soviet and later American mathematician, member of the US National Academy of Science. He has made contributions to the fields of probability and algebra. The Dynkin diagram, the Dynkin system, and Dynkin's lemma are all named after him.

Who solved the secretary problem?

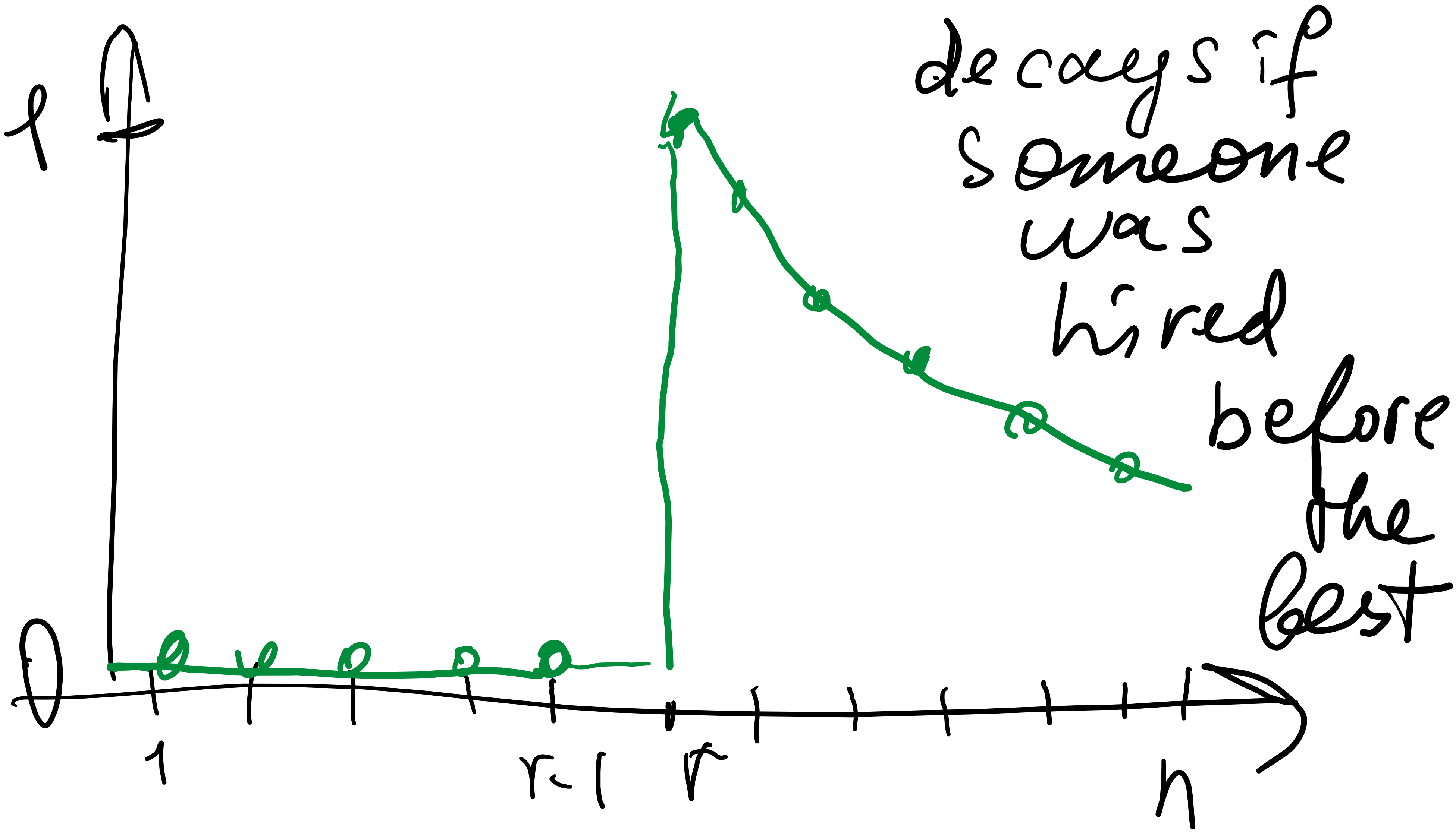
- Gardner outlined the solution in Sci Am 1960 but gave no formal proof
- Solution by Lindey was published in 1961:
  Lindey, D. V. (1961). Dynamic programming and decision theory. Appl. Statist. 10 39-51
- Dynkin's paper was published in 1963:
  Dynkin, E. B. (1963). The optimum choice of the instant for stopping a Markov process. Soviet Math. Dokl. 4 627-629
- When the celebrated German astronomer, Johannes Kepler (1571-1630), lost his first wife to cholera in 1611, he set about finding a new wife
- He spent 2 years on the process, had 11 candidates and married the 5^{th} candidate (11/e~4 so he married the first after)

Thomas S Ferguson, Statistical Science 1989, Who Solved the Secretary Problem?

What should the employer do?

- Employer does not know the distribution of the quality of applicants and has to learn it on the fly
- Algorithm: look at the first ***r-1*** applicants, remember the best among them
- Hire the first among next ***n-r+1*** applicants who is better than the best among the first ***r*** applicants
- How to choose ***r***?
- When ***r*** is too small – not enough information: the best among *r* is not very good. You are likely to hire a bad secretary
- When ***r*** is too large (e.g. ***r=n-1***) – you procrastinated for too long! You have almost all the information, but you will have to hire the last applicant who is (likely) not particularly good

Probability of hiring the best candidate if he/she has #i in the queue

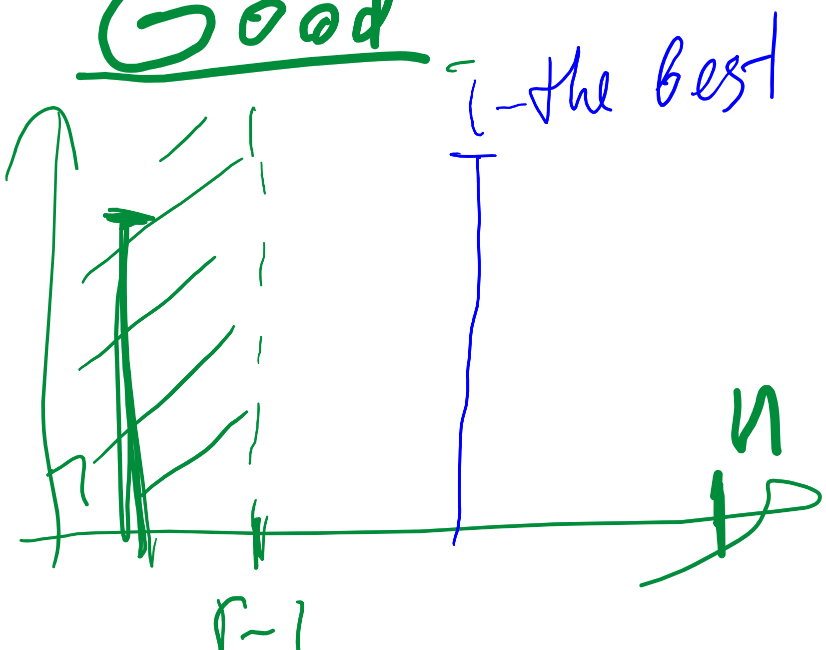


Look at $i-1$ candidates before the best

$$Prob = \frac{r-1}{i-1}$$

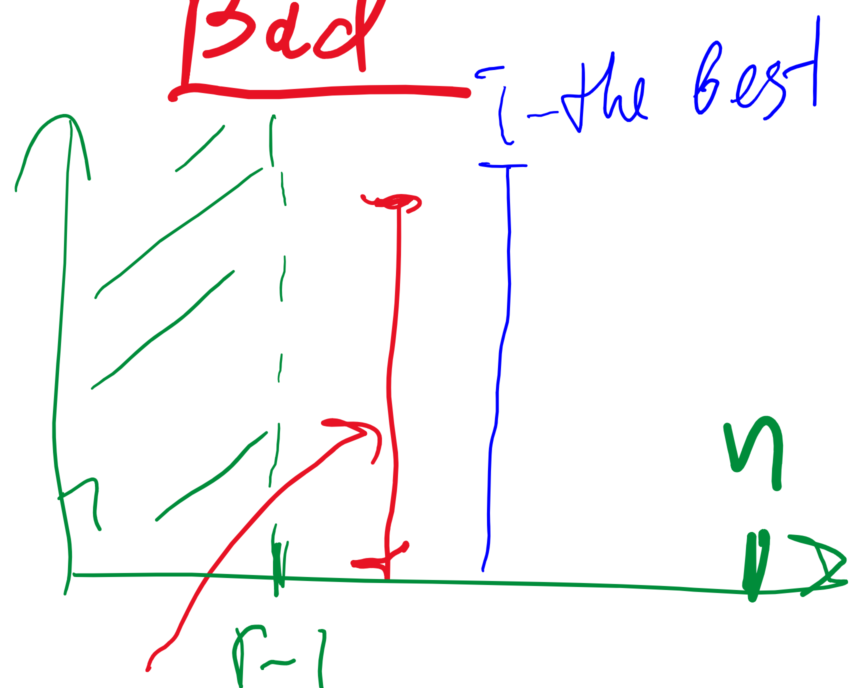
$$Prob = \frac{i-r}{i-1}$$

Good



the best among $i-1$

Bad



the best among $i-1$

$$
\begin{aligned}
P(r) &= \sum_{i=1}^{n} P\left(\text{applicant } i \text{ is selected} \cap \text{applicant } i \text{ is the best}\right) \\
&= \sum_{i=1}^{n} P\left(\text{applicant } i \text{ is selected} \middle| \text{applicant } i \text{ is the best}\right) \times P\left(\text{applicant } i \text{ is the best}\right) \\
&= \left[\sum_{i=1}^{r-1} 0 + \sum_{i=r}^{n} P\left(\begin{array}{l}\text{the best of the first } i-1 \text{ applicants} \\ \text{is in the first } r-1 \text{ applicants}\end{array} \middle| \text{applicant } i \text{ is the best}\right)\right] \times \frac{1}{n} \\
&= \sum_{i=r}^{n} \frac{r-1}{i-1} \times \frac{1}{n} \quad = \quad \frac{r-1}{n} \sum_{i=r}^{n} \frac{1}{i-1}.
\end{aligned}
$$

$$P(r) = \frac{r-1}{n}\sum_{i=r}^{n}\frac{1}{i-1}.$$

Letting *n* tend to infinity, writing x as the limit of *r*/*n*, using *t* for *i*/*n* and *dt* for 1/*n*,

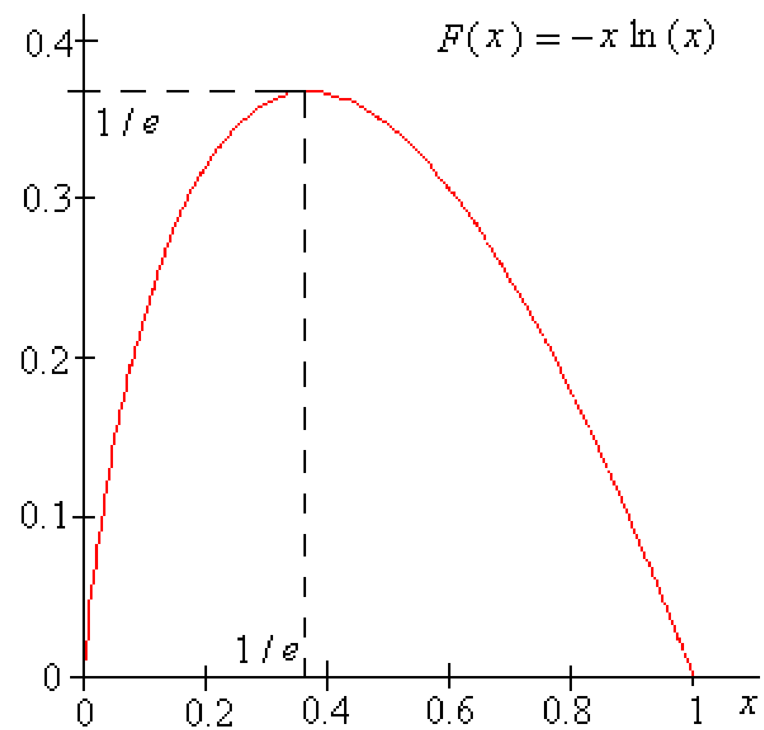
$$P(x) = x\int_{x}^{1}\frac{1}{t}\,dt = -x\ln(x)\ .$$

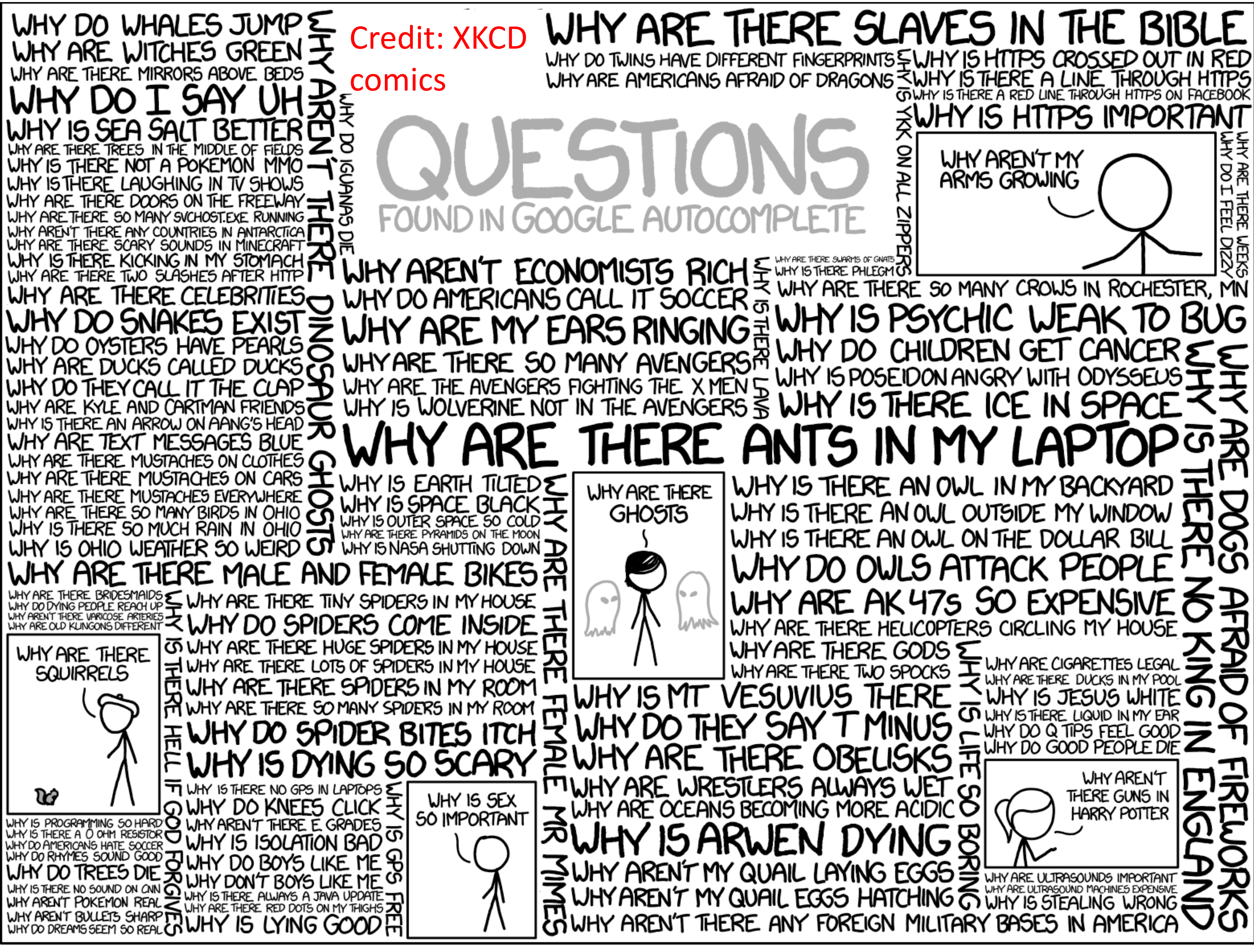
dP(x)/dx=-ln(x)-1

-ln(x)-1=0*

x=1/e=0.3679*

Probability of picking the best applicant is also *1/e=0.3679*





Credit: XKCD comics

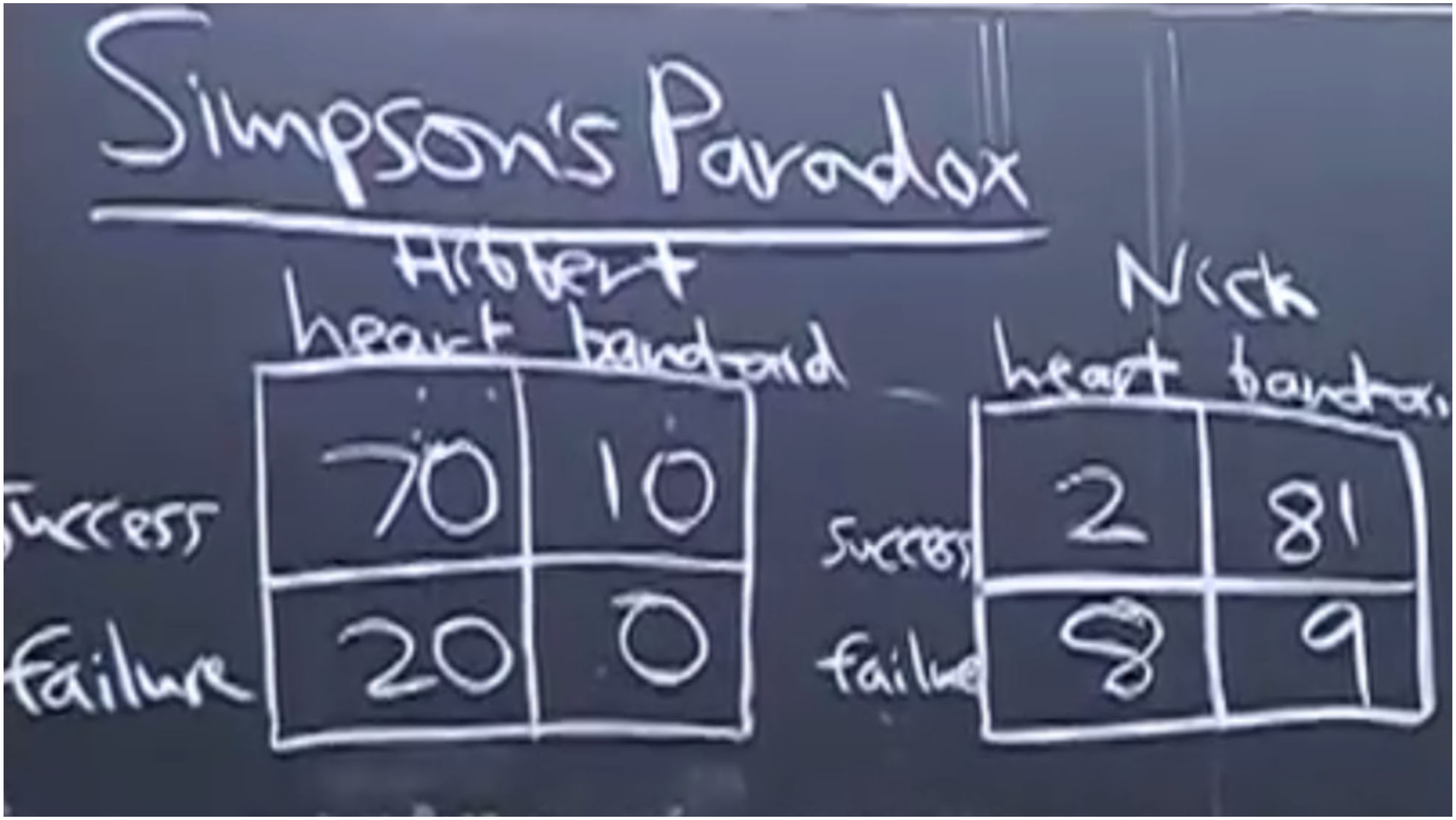
Simpson's paradox

Edward Hugh Simpson
(10 December 1922 – 5 February 2019)
was a British codebreaker, statistician
and civil servant.
"The Interpretation of Interaction
in Contingency Tables“, Journal of the Royal Statistical Society, 1951

Is it possible for one doctor to have a **higher** success rate than another doctor in **every** type of treatment he performs but to have a **lower** **overall** success rate across all treatment types?

Dr. Hibbert

Dr. Nick



Dr. Hibbert: success rate =80%

Dr. Nick: success rate =83%